SYSTEMS OF LINEAR CONGRUENCES

A. T. BUTSON AND B. M. STEWART

1. Introduction. On recent occasions papers have been presented concerned with the problem of solving a system of linear congruences. Apparently the authors were not aware that this problem was solved very neatly and completely a long time ago by H. J. S. Smith (**5**; **6**). One reason for this situation is that recent texts in the theory of numbers go only as far in the discussion of systems of congruences as one can with the most elementary tools; whereas older texts, such as the one by Stieltjes (**8**, pp. 284–377), devote so much space to the discussion of the requisite matric theory that the reader is liable to lose sight of the elegant results concerning systems of congruences. Perhaps the time has come to give a new exposition of this material, particularly since this can be done in rather short compass to an audience whose background may be assumed to include acquaintance with the invariant factors and the Smith normal form of a matrix with elements in a principal ideal ring, $\mathfrak{P}$.

In the final part of this paper we present some original work extending the discussion of systems of linear equations, and systems of linear congruences modulo an ideal, from the classical case over the rational domain to the case where the systems are over a set of integral elements, with a $\mathfrak{P}$-basis, belonging to an associative algebra. Here we assume knowledge of the Hermite normal form of a matrix with elements in a principal ideal ring.

The Classical Case

2. The problems. The coefficients, constants, and moduli in the following equations and congruences are assumed to be in a specified principal ideal ring $\mathfrak{P}$, such as the rational domain. For the system of p linear equations in n unknowns represented by

$$\sum_{i=1}^{n} x_i a_{ij} = k_j, \qquad j = 1, 2, \ldots, p,$$

we will use the matric notation

$$XA = K, \tag{1}$$

where X is 1-by-n, A is n-by-p, and K is 1-by-p.

The first problem is to determine when (1) has a solution X with elements in $\mathfrak{P}$, to find how many solutions there may be, and to give a method for actually obtaining the solution.

For the system of linear congruences represented by

$$\sum_{i=1}^{n} x_i b_{ij} \equiv g_j \pmod{m_j}, \qquad j = 1, 2, \ldots, p,$$

Received November 15, 1954.

https://doi.org/10.4153/CJM-1955-039-2 Published online by Cambridge University Press

we note that if $m = [m_1, m_2, \ldots, m_p]$ is the least common multiple of the m_j, then an equivalent system of congruences is given by

$$\sum_{i=1}^{n} x_i a_{ij} \equiv k_j \pmod{m}, \qquad j = 1, 2, \ldots, p,$$

where $a_{ij} = b_{ij}r_j$, $k_j = g_j r_j$, and $m = m_j r_j$. If we denote this system by

$$XA \equiv K \pmod{m}, \tag{2}$$

then the second problem is to answer for (2) the same three questions we have listed for (1).

3. Necessary and sufficient conditions for solving (1). There exist **(4,** Theorem 105.2**)** unimodular matrices U and V with elements in $\mathfrak{P}$, U being n-by-n and V being p-by-p, such that $UAV = E$ is in Smith normal form, with zero elements everywhere except in the main diagonal where there may appear non-zero elements $e_1, e_2, \ldots, e_r$ (which are called invariant factors and which are uniquely determined up to associates in $\mathfrak{P}$) having the property that e_i divides e_{i+1} and either $r \leqslant p \leqslant n$ or $r \leqslant n < p$.

Hence the system (1) may be replaced by the equivalent system

$$(XU^{-1})(UAV) = KV,$$

so that by setting $Y = XU^{-1}$ and $C = KV$, we arrive at

$$YE = C. \tag{3}$$

The system (3) is so simple that we can immediately conclude that necessary and sufficient conditions for its solution are as follows:

$$e_i \text{ must divide } c_i, \quad i = 1, 2, \ldots, r; \quad c_i = 0, i > r. \tag{4}$$

If we define $A' = \binom{A}{K}$ as the augmented matrix of (1), then using the conventional block notation we have

$$\begin{pmatrix} U & O \\ O & 1 \end{pmatrix} A' V = \begin{pmatrix} E \\ C \end{pmatrix},$$

so that a further transformation by unimodular matrices U' and V', where U' is $(n+1)$-by-$(n+1)$ and V' is p-by-p, will take A' into its Smith normal form, say $U'\binom{E}{C}V' = E'$, which is $(n+1)$-by-p with elements e'_i in the main diagonal. Thus depending on the relative size of n and p, the conditions (4) may be given the following form:

(5) $p \leqslant n$: $e'_i = e_i$, $i = 1, 2, \ldots, p$;

(5′) $n < p$: $e'_i = e_i$, $i = 1, 2, \ldots, n$; and $e'_{n+1} = 0$.

4. Necessary and sufficient conditions for solving (2). Since the congruence problem (2) requires the existence of elements t_j in $\mathfrak{P}$ such that

$$t_j m + \sum_{i=1}^{n} x_i a_{ij} = k_j, \qquad j = 1, 2, \ldots, p,$$

https://doi.org/10.4153/CJM-1955-039-2 Published online by Cambridge University Press

it is easy to replace the system of congruences (2) by an equivalent system of $p_1 = p$ equations in $n_1 = n + p$ unknowns, say

$$XA \ + TM = K, \tag{6}$$

where T is 1-by-p and $M = mI_p$ is a scalar p-by-p matrix.

Since $p_1 < n_1$, we apply the test (5) to the system (6). This requires us to compute the invariant factors of $\binom{A}{M}$ and $\binom{A'}{M}$. Fortunately this task is easy because of the form of M. Following an argument by Butson which is more direct than that used by Smith, we write

$$\begin{pmatrix} U & O \\ O & V^{-1} \end{pmatrix} \begin{pmatrix} A \\ M \end{pmatrix} V = \begin{pmatrix} E \\ M \end{pmatrix}.$$

Because e_i divides e_{i+1}, we see that no further arrangement of columns is necessary and that the ith invariant factor of $\binom{A}{M}$ is either (e_i, m) when $i \leqslant p \leqslant n$, or is m when $n < i \leqslant p$.

Similarly, for $\binom{A'}{M}$ the ith invariant factor is (e'_i, m) when $i \leqslant p \leqslant n$; but when $n < p$, the $(n + 1)$st invariant factor is (e'_{n+1}, m), and when $n + 1 < i \leqslant p$, the ith invariant factor is m.

Hence the test (5) shows that the necessary and sufficient conditions for the solution of (2) are as follows:

(7) $p \leqslant n$: $(e'_i, m) = (e_i, m), \quad i = 1, 2, \ldots, p$;

(7′) $n < p$: $(e'_i, m) = (e_i, m), \quad i = 1, 2, \ldots, n$; and $(e'_{n+1}, m) = m$.

We note that the final condition in (7′) may be written $e'_{n+1} \equiv 0 (\text{mod } m)$.

5. The number of solutions and their form. To determine how many solutions there are and actually to produce them, we return to (3), supposing that the necessary and sufficient conditions stated above are satisfied.

In the case of equations we see from $YE = C$, that the first r of the y's are determined uniquely by $y_i = c_i/e_i$, while the remaining $n - r$ of the y's are arbitrary. The complete solution of (1) is then given by $X = YU$ and involves $n - r$ parameters. Of course, since U and V are not unique, the complete solution may be obtained in a variety of forms, differently expressed, but actually equivalent.

In the case of congruences we consider solutions X' and X of (2) to be distinct only when $X' \not\equiv X$ (mod m), i.e., when for at least one value of i we have $x'_i \not\equiv x_i$ (mod m). We see from $YE \equiv C$ (mod m), that the first r of the y's are determined by congruences of the form $y_i e_i \equiv c_i$ (mod m). From properties of $\mathfrak{P}$ we know there are as many solutions y_i which are incongruent mod m as there are residue classes of $\mathfrak{P}$, mod (e_i, m). The remaining y's are arbitrary, so for each of these there are as many solutions incongruent mod m as there are residue classes of $\mathfrak{P}$ (mod m). The solutions of (2) are given explicitly by $X \equiv YU$ (mod m), so there are as many distinct solutions X as there are distinct solutions Y. (Moreover, we may check that $x'_i \equiv x_i$ (mod m) if and only if

$$x'_i \equiv x_i \ (\text{mod } m_j), \qquad j = 1, 2, \ldots, p;$$

https://doi.org/10.4153/CJM-1955-039-2 Published online by Cambridge University Press

so there are the same number of incongruent solutions of the original system of congruences with moduli $m_1, m_2, \ldots, m_p$.)

In particular, when $\mathfrak{P}$ is the domain of rational integers, the above considerations show that there are exactly

$$N = (e_1, m)(e_2, m) \ldots (e_r, m)\, m^{n-r}$$

distinct solutions of (2).

6. Example. We take $\mathfrak{P}$ to be the rational integers and consider the system

$$3x_1 + x_2 = 5,$$
$$5x_1 + 3x_2 = 1.$$

Using the notation of the preceding sections, we have

$$UAV = \begin{pmatrix} 0 & 1 \\ 1 & -3 \end{pmatrix}\begin{pmatrix} 3 & 5 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} 1 & 3 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} = E,$$

$$KV = (5 \quad 1)\, V = (5 \quad 14) = C,$$

$$U' \begin{pmatrix} E \\ C \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -5 & -3 & 1 \\ 10 & 7 & -2 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 4 \\ 5 & 14 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \\ 0 & 0 \end{pmatrix} = E'.$$

Since $4 = e_2 \neq e'_2 = 2$, it follows from (5) that the system has no solution in rational integers.

Considering the same system mod m, we see from (7) that we must have $(1, m) = (1, m)$ and $(4, m) = (2, m)$. If $m \equiv 0 \pmod 4$, there are no solutions; if $m \equiv 1$ or $3 \pmod 4$, there will be $N = 1$ solutions; and if $m \equiv 2 \pmod 4$, there will be $N = 2$ solutions.

Thus if $m = 10$, we solve $y_1 \equiv 5$, $4y_2 \equiv 14$, for $y_1 \equiv 5$, $y_2 \equiv 1$; and $y_1 \equiv 5$, $y_2 \equiv 6$. Then from $X \equiv YU$, we compute $x_1 \equiv 1$, $x_2 \equiv 2$; and $x_1 \equiv 6$, $x_2 \equiv 7$, respectively.

The Case of Integral Elements of an Algebra

7. Sets of integral elements. Let $\mathfrak{A}$ be an associative algebra defined over a field $\mathfrak{F}$ and possessing a modulus ϵ. Suppose that $\mathfrak{F}$ contains the principal ideal ring $\mathfrak{P}$. Each element of a set $\mathfrak{O}$ of elements of $\mathfrak{A}$ will be called an integral element if the set has the following three properties:

U (unity): the set contains the modulus ϵ;

C (closure): the set is closed under addition, subtraction, and multiplication;

B (finite basis): the set contains elements $\epsilon_1 = \epsilon, \epsilon_2, \ldots, \epsilon_k$ such that every element of the set is expressed uniquely in the form

$$\sum a_i \epsilon_i$$

where each a_i is in $\mathfrak{P}$.

As an example we may take $k = 1$ and obtain as $\mathfrak{O}$ the ring $\mathfrak{P}$ itself. Again when $\mathfrak{F}$ is the rational field and $\mathfrak{P}$ the rational domain, we see that $\mathfrak{O}$ is a set of integral elements (but not necessarily a maximal set) in the sense of Dickson **(1**, Chapter X**)**. See also **(3)**.

8. The regular representations. If $\alpha = \sum a_i \epsilon_i$ is any element of $\mathfrak{Q}$, by properties C and B there must exist elements r_{tj} and s_{jt} of $\mathfrak{P}$ such that

$$\alpha \epsilon_j = (\sum a_i \epsilon_i) \epsilon_j = \sum r_{tj} \epsilon_t, \qquad j = 1, 2, \ldots, k;$$
$$\epsilon_j \alpha = \epsilon_j(\sum a_i \epsilon_i) = \sum s_{jt} \epsilon_t, \qquad j = 1, 2, \ldots, k.$$

Hence with each α in $\mathfrak{Q}$ there are associated k-by-k matrices $R(\alpha) = (r_{tj})$ and $S(\alpha) = (s_{jt})$ with elements in $\mathfrak{P}$.

If $\mathbf{E}$ indicates the 1-by-k matrix with elements $\epsilon_1, \epsilon_2, \ldots, \epsilon_k$, then in matric notation we have

$$\alpha \mathbf{E} = \mathbf{E} R(\alpha), \quad \mathbf{E}^T \alpha = S(\alpha)\mathbf{E}^T. \tag{8}$$

Since $\epsilon_1 = \epsilon$, the first column of $R(\alpha)$ and the first row of $S(\alpha)$ consist of precisely the elements $a_1, a_2, \ldots, a_k$. Then from property B it follows that the correspondences defined by (8) are both one-to-one. Moreover, it is easily shown that the correspondences are preserved under the addition and multiplication operations of each system. Hence the matrices $R(\alpha)$ and the matrices $S(\alpha)$ provide isomorphic representations for $\mathfrak{Q}$, well known, respectively, as the first and second regular representations.

If α and β are in $\mathfrak{Q}$, we may use (8) and the fact that elements of $\mathfrak{P}$ commute with elements of $\mathfrak{Q}$ to write

$$\begin{aligned} R^T(\alpha)S(\beta)\mathbf{E}^T &= R^T(\alpha)\mathbf{E}^T \beta = \alpha I\mathbf{E}^T \beta \\ &= \alpha IS(\beta)\mathbf{E}^T = S(\beta)\alpha I\mathbf{E}^T = S(\beta)R^T(\alpha)\mathbf{E}^T; \end{aligned}$$

then from property B, it follows that

$$R^T(\alpha)S(\beta) = S(\beta)R^T(\alpha).$$

In particular, letting $A = (a_1, a_2, \ldots, a_k)$ and $B = (b_1, b_2, \ldots, b_k)$ be the first rows of $R^T(\alpha)$ and $S(\beta)$, respectively, we obtain the useful relation

$$AS(\beta) = BR^T(\alpha). \tag{9}$$

9. Systems of linear equations over $\mathfrak{Q}$. We consider the following system of p linear equations in n unknowns:

$$\sum_{i=1}^{n} \alpha_{ij}\, \chi_i\, \beta_{ij} = \gamma_j, \qquad j = 1, 2, \ldots, p, \tag{10}$$

where the α_{ij}, β_{ij}, and γ_j are given elements of $\mathfrak{Q}$. Since $\mathfrak{Q}$ is not necessarily commutative, note that coefficients are allowed on both sides of the unknowns. We are concerned to establish necessary and sufficient conditions that (10) have solutions $\chi_1, \chi_2, \ldots, \chi_n$ which are in $\mathfrak{Q}$.

If we assume that such solutions exist we may write $\chi_i = \sum x_{ij} \epsilon_j$, where the x_{ij} are in $\mathfrak{P}$, and define $X_i = (x_{i1}, \ldots, x_{ik})$. Supposing $\gamma_j = \sum c_{ji} \epsilon_i$, we define $C_j = (c_{j1}, \ldots, c_{jk})$. We define A_{ij} to be the first row of $S(\alpha_{ij})$. Then (8), (9) and (10) imply that

$$C_j \mathbf{E}^T = \sum A_{ij} \mathbf{E}^T \chi_i \beta_{ij} = \sum A_{ij} S(\chi_i) \mathbf{E}^T \beta_{ij}$$
$$= \sum X_i R^T(\alpha_{ij}) S(\beta_{ij}) \mathbf{E}^T.$$

Hence property B implies that

$$C_j = \sum_{i=1}^{n} X_i R^T(\alpha_{ij}) S(\beta_{ij}), \qquad j = 1, 2, \ldots, p.$$

We set

$$C = (c_{11}, \ldots, c_{1k}; c_{21}, \ldots, c_{2k}; c_{p1}, \ldots, c_{pk}),$$
$$X = (x_{11}, \ldots, x_{1k}; x_{21}, \ldots, x_{2k}; \ldots; x_{n1}, \ldots, x_{nk}),$$

and $A = (R^T(\alpha_{ij}) S(\beta_{ij}))$ where C is 1-by-pk, X is 1-by-nk, and the "enlarged coefficient matrix" A is nk-by-pk made up of k-by-k blocks of which the one in the ij-position is $R^T(\alpha_{ij}) S(\beta_{ij})$. Then the equations obtained above may be written as the single matric equation

$$XA = C. \tag{11}$$

Except for the size of the matrices involved, (11) is precisely a system of the classical type (1) with kp equations in kn unknowns, with the elements involved all in $\mathfrak{P}$.

Conversely, if (11) has a solution X in $\mathfrak{P}$, we can retrace the steps above to obtain in $\mathfrak{O}$ a solution of (10). Moreover, by property B, distinct solutions of (11) lead to distinct solutions of (10).

Thus the problem of solving (10) in $\mathfrak{O}$ has been shown equivalent to solving (11) in $\mathfrak{P}$. Referring to (5) and (5′) we can assert that if $e_1, e_2, \ldots$ are the invariant factors of A and if $e'_1, e'_2, \ldots$ are the invariant factors of the augmented matrix $\binom{A}{C}$, then necessary and sufficient conditions that the system (10) have a solution are that

(12) $\quad p \leqslant n$: $\quad e'_i = e_i, \quad i = 1, 2, \ldots, kp$;
(12′) $\quad n < p$: $\quad e'_i = e_i, \quad i = 1, 2, \ldots, kn$; and $e_{kn+1}' = 0$.

Determining the number of solutions and the most general solution proceeds along the lines given in §5. In these matters it is worth a word of caution that the rank of A need not be a multiple of k.

We note, thanks to the referee, that one type of matric equation, well known in the literature, is included in the above discussion. For if the algebra $\mathfrak{A}$ is a total matric algebra of order $k = n^2$, having the natural basis of elements ϵ_{ij}, where ϵ_{ij} is an n-by-n matrix with 1 in the ij position and zeros elsewhere, so that the multiplication table is

$$\epsilon_{ij}\, \epsilon_{rs} = \delta_{jr}\, \epsilon_{is},$$

and if $\mathbf{E} = (\epsilon_{11}, \ldots, \epsilon_{1n}; \epsilon_{21}, \ldots, \epsilon_{2n}; \ldots; \epsilon_{n1}, \ldots, \epsilon_{nn})$, then the typical element $\beta = \sum b_{ij}\, \epsilon_{ij}$, which we ordinarily represent as $B = (b_{ij})$, has the regular representations

$$R(\beta) = I \cdot \times B, \quad S(\beta) = B \cdot \times I,$$

where $M \cdot \times B$ indicates the direct product matrix which is n^2-by-n^2 and whose ij block is Mb_{ij}. Then a linear equation like $\alpha\, \chi\, \beta = \gamma$ is replaced, according to the theory above for passing from (10) to (11), by an equation $X'\, D' = C'$, where

$$D' = R^T(\alpha)\, S(\beta) = (I \cdot \times A)^T (B \cdot \times I) = B \cdot \times A^T$$

and X' and C' are 1-by-n^2, obtained from $X = (x_{ij})$ and $C = (c_{ij})$, respectively, by taking row blocks. The general linear equation in one unknown $\sum \alpha_i\, \chi\, \beta_i = \gamma$ may be treated in the same manner, the enlarged coefficient matrix being $D'' = \sum B_i \cdot \times A_i{}^T$. This is the *nivellateur* studied by Sylvester **(9)**, however, only for the case $\mathfrak{P} = \mathfrak{F}$.

Similarly, the system of equations (10) may be generalized to allow each unknown to appear in a finite number of summands in each equation; the technique for passing to (11) remains the same, except each component block of the enlarged coefficient matrix will now be a sum of matrices of the type $R^T(\alpha_{ij})\, S(\beta_{ij})$.

10. Minimal bases for ideals in $\mathfrak{Q}$. In the usual manner the left ideal $\mathfrak{M}$ generated by $\zeta_1, \zeta_2, \ldots, \zeta_t$, a given set of elements of $\mathfrak{Q}$, is defined to be the set of elements

$$\sum_{i=1}^{t} \nu_i\, \zeta_i$$

obtained by allowing the left-multipliers ν_i to vary independently over all of $\mathfrak{Q}$. A minimal basis for the ideal $\mathfrak{M}$ is by definition a set of elements $\mu_1, \mu_2, \ldots, \mu_s$ such that an element of $\mathfrak{Q}$ is in the ideal $\mathfrak{M}$ if and only if it can be represented in the form

$$\sum_{i=1}^{s} c_i\, \mu_i,$$

where the c_i are in $\mathfrak{P}$; and this representation is to be unique.

An argument by MacDuffee **(2)** shows that if H is the uniquely determined left-Hermite form of the matrix

$$S = \begin{pmatrix} S(\zeta_1) \\ \cdots \\ S(\zeta_t) \end{pmatrix},$$

then the non-zero rows H_i of H determine a minimal basis for $\mathfrak{M}$, having $s \leqslant k$, by the relation $\mu_i = H_i\, \mathbf{E}^T$. The notation which we have been using makes it simple to reproduce the proof.

Let U be a unimodular matrix, kt-by-kt, having elements in $\mathfrak{P}$, such that $US = H$. Let $V = U^{-1}$, so that $S = VH$. If the ith row of U is divided into 1-by-k blocks U_{ij}, then

$$\mu_i = H_i\, \mathbf{E}^T = \sum U_{ij}\, S(\zeta_j)\, \mathbf{E}^T = \sum U_{ij}\, \mathbf{E}^T\, \zeta_j = \sum \nu_{ij}\, \zeta_j,$$

where $\nu_{ij} = U_{ij}\, \mathbf{E}^T$ is in $\mathfrak{Q}$, hence μ_i is in the ideal $\mathfrak{M}$, and so are all $\sum c_i\, \mu_i$.

Conversely, given any element $\nu = \sum \nu_i \zeta_i$ in the ideal, we have $\nu_i = \sum n_{ij} \epsilon_j$ and if we define $N_i = (n_{i1}, \ldots, n_{ik})$ we can write $\nu_i = N_i \mathbf{E}^T$. Hence

$$\nu = \sum N_i \mathbf{E}^T \zeta_i = \sum N_i S(\zeta_i) \mathbf{E}^T = NS\mathbf{E}^T = NVH\mathbf{E}^T = \sum c_i \mu_i$$

where N is 1-by-kt, made up of the 1-by-k blocks N_i, and where c_i is the element in the ith column of NV. Since c_i is in $\mathfrak{P}$, a representation of the desired type for ν has been found. The uniqueness of the representation follows from the independence of the non-zero rows in the canonical left-Hermite form H.

In an analogous way we define the right-ideal generated by $\zeta_1, \zeta_2, \ldots, \zeta_t$ to be the set of elements $\sum \zeta_i \eta_i$ obtained by allowing the right-multipliers η_i to vary independently over $\mathfrak{O}$. In this case a minimal basis can be found by computing the left-Hermite form D of the matrix R which is kt-by-k with its ith k-by-k block being $R^T(\zeta_i)$; for if $D_1, \ldots, D_r$ are the non-zero rows of D, necessarily with $r \leqslant k$, then the elements $\delta_j = D_j \mathbf{E}^T$ serve as a minimal basis.

By combining these observations we can find a minimal basis for the two-sided ideal generated by $\zeta_1, \zeta_2, \ldots, \zeta_t$ whose typical element is

$$\alpha = \sum_{i=1}^{t} \sum_{j=1}^{q_i} \nu_{ij} \zeta_i \eta_{ij},$$

where the q_i are all finite. For we may first compute a minimal basis $\mu_1, \mu_2, \ldots, \mu_s$ for the left ideal generated by $\zeta_1, \zeta_2, \ldots, \zeta_t$ and replace each $\nu_{ij} \zeta_i$ by $\sum c_{ijm} \mu_m$. Then

$$\alpha = \sum_{m=1}^{s} \mu_m \eta_m, \text{ where } \eta_m = \sum_{i=1}^{t} \sum_{j=1}^{q_i} c_{ijm} \eta_{ij}.$$

Hence if, secondly, we compute a minimal basis $\delta_1, \delta_2, \ldots, \delta_r$ for the right ideal generated by $\mu_1, \mu_2, \ldots, \mu_s$, we shall have arrived at a suitable minimal basis $\delta_1, \delta_2, \ldots, \delta_r$ for the two-sided ideal generated by $\zeta_1, \zeta_2, \ldots, \zeta_t$.

However, not every matrix H in left-Hermite form represents a minimal basis for an ideal of $\mathfrak{O}$ (**2**, p. 76).

When α and β are in $\mathfrak{O}$, by the notation $\alpha \equiv \beta \pmod{\mathfrak{M}}$ we mean that $\alpha - \beta$ is in the ideal $\mathfrak{M}$ and we say that α and β are in the same residue class mod $\mathfrak{M}$. For the sequel it is important to notice that, in general, it is only when the ideal $\mathfrak{M}$ is two-sided that multiplication of residue classes mod $\mathfrak{M}$ is well defined.

11. Systems of linear congruences modulo ideals. Over $\mathfrak{O}$ we consider the following system of p linear congruences, modulo ideals of $\mathfrak{O}$, in n unknowns:

$$(13) \qquad \sum_{i=1}^{n} \alpha_{ij} \chi_i \beta_{ij} \equiv \gamma_j \pmod{\mathfrak{M}_j}, \qquad j = 1, 2, \ldots, p.$$

We shall assume as explained in §10 that for the ideal $\mathfrak{M}_j$, whether it be left, right, or two-sided, a minimal basis of s_j elements has been found, say

$$\mu_{1j}, \mu_{2j}, \ldots, \mu_{s_j j},$$

given by $\mu_{ij} = H_{ij} \mathbf{E}^T$ where the H_{ij} are non-zero rows of a left-Hermite

k-by-k matrix. We let H_j be the s_j-by-k matrix with rows H_{ij}, so that H_j is what is called an echelon row form.

The system (13) is then equivalent to the system

$$\sum_{i=1}^{n} \alpha_{ij}\, \chi_i\, \beta_{ij} + \sum_{i=1}^{s_j} t_{ij}\, \mu_{ij} = \gamma_j, \qquad j = 1, 2, \ldots, p,$$

where the unknowns t_{ij} are in $\mathfrak{P}$. Following the same development and using the same notation as in §9, we may show that solving (13) in $\mathfrak{Q}$ is equivalent to solving in $\mathfrak{P}$ the following system:

$$(14) \qquad (X\ T)\begin{pmatrix} A \\ H \end{pmatrix} = C,$$

where

$$T = (t_{11}, \ldots, t_{s_1 1}; t_{12}, \ldots, t_{s_2 2}; \ldots ; t_{1p}, \ldots, t_{s_p p})$$

and H is the direct sum $H = H_1 \dotplus H_2 \dotplus \ldots \dotplus H_p$.

The number of unknowns in (14) is $nk + s$, where $s = \sum s_j$, and the number of equations is pk.

If $pk \leqslant nk + s$, we apply (5) to obtain the conditions

$$(15) \qquad e_i \begin{pmatrix} A \\ C \\ H \end{pmatrix} = e_i \begin{pmatrix} A \\ H \end{pmatrix}, \qquad i = 1, 2, \ldots, pk.$$

If $nk + s < pk$, we apply (5′) to obtain the conditions

$$(15') \qquad e_i \begin{pmatrix} A \\ C \\ H \end{pmatrix} = \begin{cases} e_i \begin{pmatrix} A \\ H \end{pmatrix}, & i = 1, 2, \ldots, nk + s, \\ 0 & , i = nk + s + 1. \end{cases}$$

Thus (15) and (15′) represent necessary and sufficient conditions for the solution of (13).

If the solution is obtained as in §5, starting from (14), unnecessary parameters may be noticed. We have made a further study of (14) in which the Smith forms D_j of H_j play a part, as well as the least common multiple m of the elements of all the D_j. This method seems of some interest because of avoiding unnecessary parameters, but the alternative set of conditions which is obtained lacks the directness of (15) and (15′). This further study emphasized the need of care in the definitions of congruent solutions.

Suppose that $\chi_1, \chi_2, \ldots, \chi_n$ is a solution of (13). If all the ideals $\mathfrak{M}_j$ are two-sided and if

$$(16) \qquad \chi'_i \equiv \chi_i \pmod{\mathfrak{M}_j}, \qquad i = 1, 2, \ldots, n; j = 1, 2, \ldots, p;$$

then $\chi'_1, \chi'_2, \ldots, \chi'_n$ is also a solution of (16). But if one or more of the ideals $\mathfrak{M}_j$ is one-sided, (16) is no longer sufficient to guarantee that $\chi'_1, \chi'_2, \ldots, \chi'_n$ is a solution of (13). Having given these words of caution, we now define sets of solutions of (13) which satisfy (16) to be congruent sets.

In matric form (16) may be written

$$X'_i - X_i = W_{ij} H_j, \qquad i = 1, 2, \ldots, n; j = 1, 2, \ldots, p.$$

Hence $X'_i - X_i$ is a common left multiple of the H_j. By repeated application of the method described in **(7)**, there is a constructive way of finding M the least common left multiple of $H_1, H_2, \ldots, H_p$. Then (16) is equivalent to the existence of Q_i with elements in $\mathfrak{P}$ such that

$$X'_i - X_i = Q_i M, \qquad i = 1, 2, \ldots, n;$$

and hence to the single matric equation

$$X' - X = QM^* \tag{17}$$

where $M^* = M \dotplus M \dotplus \ldots \dotplus M$ with n summands. Hence we may apply (5) and (5′) to obtain conditions equivalent to (16) expressed in terms of the invariant factors of M^* and

$$\begin{pmatrix} M^* \\ X' - X \end{pmatrix}.$$

12. Example. Letting $\mathfrak{F}$ be the rational field and $\mathfrak{P}$ the rational domain, we consider the algebra $\mathfrak{A}$ having as a basis $\epsilon_1 = \epsilon, \epsilon_2, \epsilon_3$ with $\epsilon_2 \epsilon_2 = \epsilon_2, \epsilon_3 \epsilon_2 = \epsilon_3$, and $\epsilon_2 \epsilon_3 = \epsilon_3 \epsilon_3 = 0$. If we take as $\mathfrak{O}$ the set of all $\alpha = a\,\epsilon_1 + b\,\epsilon_2 + c\,\epsilon_3$ where a, b, c are in $\mathfrak{P}$, we have a set of integral elements with the basis $\epsilon_1, \epsilon_2, \epsilon_3$. Using (8) we find

$$R^T(\alpha) = \begin{pmatrix} a & b & c \\ 0 & a+b & c \\ 0 & 0 & a \end{pmatrix}, \qquad S(\alpha) = \begin{pmatrix} a & b & c \\ 0 & a+b & 0 \\ 0 & 0 & a+b \end{pmatrix}.$$

Taking $\alpha = (3, 3, 1)\,\mathbf{E}^T$, $\beta = (1, 5, 2)\,\mathbf{E}^T$, $\gamma = (0, 0, 2)\,\mathbf{E}^T$, $\zeta = (6, 2, 12)\,\mathbf{E}^T$, we will study

$$\alpha \chi \beta = \gamma, \tag{18}$$

$$\alpha \chi \beta \equiv \gamma \ (\mathrm{mod}\ (\zeta]), \tag{19}$$

$$\alpha \chi \beta \equiv \gamma \ (\mathrm{mod}\ [\zeta)), \tag{20}$$

where $(\zeta]$ and $[\zeta)$ indicate, respectively, the left and right ideals generated by ζ. First we compute

$$A = R^T(\alpha)S(\beta) = \begin{pmatrix} 3 & 3 & 1 \\ 0 & 6 & 1 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 5 & 2 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{pmatrix} = \begin{pmatrix} 3 & 33 & 12 \\ 0 & 36 & 6 \\ 0 & 0 & 18 \end{pmatrix}, \quad E = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 108 \end{pmatrix}.$$

When we find that $A' = \binom{A}{C}$ has $e'_1 = 1$, $e'_2 = 6$, $e'_3 = 108$, it follows from (12) that there is no solution to (18).

Since $S(\zeta)$ has the left-Hermite form

$$L = \begin{pmatrix} 24 & 0 & 0 \\ 12 & 4 & 0 \\ 6 & 2 & 4 \end{pmatrix},$$

we find that $\binom{A}{L}$ has $e_1 = 1$, $e_2 = 2$, $e_3 = 12$, and that $\binom{A'}{L}$ has $e'_1 = 1$, $e'_2 = 2$, $e'_3 = 12$, so by (15) there is a solution of (19).

Since $R^T(\zeta)$ has the left-Hermite form

$$R = \begin{pmatrix} 24 & 0 & 0 \\ 6 & 2 & 0 \\ 0 & 0 & 6 \end{pmatrix},$$

we find that $\binom{A}{R}$ has $e_1 = 1$, $e_2 = 6$, $e_3 = 12$, but that $\binom{A'}{R}$ has $e'_1 = 1$, $e'_2 = 2$, $e'_3 = 12$, so by (15) there is no solution to (20).

When we carry through the actual solution of (19) we find that the most general solution involves three parameters:

$$x_1 = -24z_1 + 2z_3, \quad x_2 = 22z_1 - 2z_3, \quad x_3 = 1 + 78z_1 + 4z_2 - 12z_3.$$

When we apply (17) we find that pairs of the solutions are congruent mod $(\zeta]$, if and only if

$$z'_1 \equiv z_1 (\text{mod } 4), \quad z'_2 \equiv z_2 (\text{mod } 2), \quad z'_3 \equiv z_3 (\text{mod } 3).$$

References

1. L. E. Dickson, *Algebras and their arithmetics*, (Chicago, 1923).
2. C. C. MacDuffee, *An introduction to the theory of ideals in linear associative algebras*, Trans. Amer. Math. Soc., *31* (1929), 71–90.
3. ———, *Modules and ideals in a Frobenius algebra*, Monatsh. Math., *48* (1939), 293–313.
4. ———, *An introduction to abstract algebra* (New York, 1940).
5. H. J. S. Smith, *On systems of linear indeterminate equations and congruences*, Phil. Trans. Royal Soc. London, A *151* (1861), 293–326.
6. ———, *On the arithmetical invariants of a rectangular matrix of which the constituents are integral numbers*, Proc. London Math. Soc., *4* (1873), 236–249.
7. B. M. Stewart, *A note on least common left multiples*, Bull. Amer. Math. Soc., *55* (1949), 587–591.
8. T. J. Stieltjes, *Oeuvres complètes*, II (Math. Soc. Amsterdam, 1918).
9. J. J. Sylvester, C.R. Acad. Sci., Paris, *99* (1884), 117–118, 409–412, 432–436, 527–529.

Michigan State College